Third-order density-functional tight-binding combined with the fragment molecular orbital method

AUTHOR(S):

Nishimoto, Yoshio; Fedorov, Dmitri G.; Irle, Stephan

Third-order density-functional tight-binding combined with the fragment molecular orbital method

Yoshio Nishimoto[a,b,*], Dmitri G. Fedorov[c,**], Stephan Irle[a,d]

[a] *Department of Chemistry, Nagoya University, Nagoya 466-8602, Japan*
[b] *Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103, Japan*
[c] *Nanosystem Research Institute (NRI), National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8565, Japan*
[d] *Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Nagoya 466-8602, Japan*

Abstract

We developed the energy and its analytic gradient for the self-consistent-charge density-functional tight-binding method with the third-order expansion (DFTB3) combined with the fragment molecular orbital (FMO) method, FMO-DFTB3. FMO-DFTB3 reproduced full DFTB3 relative stabilities and the optimized structures of three polyalanine isomers. FMO-DFTB3 was applied to optimize a nano flake of cellulose Iβ, consisting of 10,944 atoms, and a good agreement with the experimental structure was obtained. For a cellulose sheet containing 1,368 atoms, FMO-DFTB3 was 43.5 times faster than full DFTB3. The binding between sheets and chains in cellulose was elucidated, and two dispersion models were compared.

Keywords: fragment molecular orbital, density-functional tight-binding, third-order correction

1. Introduction

Recent methodological and algorithmic developments in computational sciences have produced a variety of efficient quantum-mechanical (QM) and molecular-mechanical (MM) approaches. QM calculations, for example, with

*Corresponding author. Tel: +81 (0)75 711-7894, E-mail address: nishimoto.yoshio@fukui.kyoto-u.ac.jp.
**Corresponding author. Tel: +81 (0)29 851-5426, e. mail d.g.fedorov@aist.go.jp.

quantum Monte-Carlo [1, 2] or post Hartree-Fock [3, 4] methods can accurately predict molecular and electronic properties; however, the computational cost is high. For instance, for coupled cluster with singles, doubles and perturbative triples (CCSD(T)) [4], also known as the "golden standard" method, the cost scales as N^7 for the system size N. On the other hand, using MM, large scale molecular dynamics simulations for 100 ns have been reported for a 10 million atomic system [5]. To reduce the cost of QM methods, a number of linear-scaling [6] as well as fragmentation methods [7, 8, 9, 10, 11, 12, 13, 14] have been developed.

The density-functional tight-binding (DFTB) method [15, 16, 17, 18] is based on an expansion of the electron density in a Taylor expansion around a reference density (usually, atomic), truncated at the linear (DFTB1), quadratic (DFTB2) or cubic (DFTB3) terms. DFTB1 is the oldest approach, which in general works well for nonpolar systems; DFTB2 adds charge transfer and polarization effects, and DFTB3 improves on the accuracy of the charge distribution, which is, in particular, important for hydrogen bonding and polar systems [19, 20, 21]. It has been shown DFTB3 is systematically better than DFTB2 [21, 22, 18, 23, 24].

The fragment molecular orbital (FMO) method [25, 26, 27, 28] is a fragment-based approach, in which a molecular system is divided into fragments, and the total properties such as the energy and its gradient are obtained from the calculations of fragments (monomers) and their pairs (dimers). Previously, we developed FMO-based DFTB2, FMO-DFTB2 [29], and showed its accuracy and efficiency.

In this work, we formulate the energy and its analytic gradient for FMO-DFTB3 and discuss the role of the damping [19, 21, 30] of the Coulomb interaction used in DFTB2 and DFTB3, in particular to improve the accuracy in hydrogen-bonded systems. The developed method is tested on three isomers of polyalanine. An analysis of the interactions in a different cellulose, Iα, has been previously conducted [31] using FMO, and here we focus on the structure optimization of the Iβ form.

2. Methodology

2.1. Overview of DFTB3

The total energy of DFTB3, E^{DFTB3}, is given by (more details can be found elsewhere [18, 19, 20, 21, 30])

$$E^{\mathrm{DFTB3}} = \sum_i n_i \sum_{\mu\nu} c_{\mu i} c_{\nu i} H^0_{\mu\nu} + \sum_{A>B} E^{\mathrm{rep}}_{AB} + \frac{1}{2}\sum_{A,B} \gamma_{AB}\Delta q_A \Delta q_B + \frac{1}{6}\sum_{A,B}(\Gamma_{AB}\Delta q_A + \Gamma_{BA}\Delta q_B)\Delta q_A \Delta q_B \ . \quad (1)$$

where μ and ν are atomic orbitals (AOs), A and B are atoms, and n_i is the occupation number of molecular orbital (MO) i, which is defined by the expansion coefficients $c_{\mu i}$. The repulsive energy E^{rep}_{AB} and the non-perturbed Hamiltonian $H^0_{\mu\nu}$ are precomputed and tabulated for model systems. γ_{AB} and Γ_{AB} are distance dependent functions used to compute the Coulombic interaction, and both of them also depend on the chemical hardness [32] of atoms. Γ_{AB} found only in DFTB3 depends also on the derivative of the Hubbard parameter with respect to the atomic Mulliken charges [30]. DFTB1 and DFTB2 energies may be obtained from Eq. (1) by neglecting the last 2 and 1 sums, respectively.

The atomic Mulliken charges Δq_A are obtained as the difference between the Mulliken population on atom A and a reference value q^0_A, $\Delta q_A = q_A - q^0_A$, where

$$q_A = \frac{1}{2}\sum_i n_i \sum_{\mu\in A}\sum_{\nu}(c_{\mu i} c_{\nu i} S_{\mu\nu} + c_{\nu i} c_{\mu i} S_{\nu\mu}) \ , \quad (2)$$

MOs are obtained by solving $\mathbf{HC} = \varepsilon\mathbf{SC}$ in self-consistent charge (SCC) calculations. $S_{\mu\nu}$ are AO overlaps and

$$H_{\mu\nu} = H^0_{\mu\nu} + S_{\mu\nu}\sum_C \left\{ \frac{1}{2}(\gamma_{AC} + \gamma_{BC}) + \frac{1}{3}(\Gamma_{AC}\Delta q_A + \Gamma_{BC}\Delta q_B) + \frac{1}{6}(\Gamma_{CA} + \Gamma_{CB})\Delta q_C \right\} \Delta q_C \ , \quad (3)$$

where $\mu \in A$ and $\nu \in B$.

The nuclear gradient of the DFTB3 energy (Eq. (1)) with respect to the x coordinate of atom α, $R_{\alpha x}$ (thoughout, x in the derivatives can be replaced

by y or z), is

$$\begin{aligned}\frac{\partial E^{\mathrm{DFTB3}}}{\partial R_{\alpha x}} = &\sum_{A\neq\alpha}\sum_{i} n_i \sum_{\mu\in A}\sum_{\nu\in\alpha} c_{\mu i}c_{\nu i}\left[2\frac{\partial H^0_{\mu\nu}}{\partial R_{\alpha x}} - 2\varepsilon_i\frac{\partial S_{\mu\nu}}{\partial R_{\alpha x}}\right.\\ &+\frac{\partial S_{\mu\nu}}{\partial R_{\alpha x}}\sum_C\left\{(\gamma_{AC}+\gamma_{\alpha C})\right.\\ &\left.\left.+\frac{1}{3}(2\Gamma_{AC}\Delta q_A+\Gamma_{CA}\Delta q_C+2\Gamma_{\alpha C}\Delta q_\alpha+\Gamma_{C\alpha}\Delta q_C)\right\}\Delta q_C\right]\\ &+\Delta q_\alpha\sum_{A\neq\alpha}\Delta q_A\frac{\partial\gamma_{A\alpha}}{\partial R_{\alpha x}}\\ &+\frac{1}{3}\Delta q_\alpha\sum_{A\neq\alpha}\Delta q_A\left(\Delta q_A\frac{\partial\Gamma_{A\alpha}}{\partial R_{\alpha x}}+\Delta q_\alpha\frac{\partial\Gamma_{\alpha A}}{\partial R_{\alpha x}}\right)\\ &+\sum_{A\neq\alpha}\frac{\partial E^{\mathrm{rep}}_{A\alpha}}{\partial R_{\alpha x}}\ .\end{aligned} \tag{4}$$

2.2. Formulation of FMO-DFTB3

The total energy E in the two-body FMO expansion (FMO2) [29, 33] is given by

$$E=\sum_I^N E'_I+\sum_{I>J}^N(E'_{IJ}-E'_I-E'_J)+\sum_{I>J}^N\Delta E^V_{IJ} \tag{5}$$

where N is the number of fragments, and the E'_X is the internal energy of fragments ($X=I$) or their pairs ($X=IJ$), called monomers and dimers, respectively. The internal energy in DFTB3 is defined as

$$\begin{aligned}E'_X=&\sum_{i\in X}n_i^X\sum_{\mu\nu\in X}c^X_{\mu i}c^X_{\nu i}H^{0,X}_{\mu\nu}+\sum_{A>B}E^{\mathrm{rep}}_{AB}\\ &+\frac{1}{2}\sum_{A,B\in X}\gamma_{AB}\Delta q^X_A\Delta q^X_B+\frac{1}{6}\sum_{A,B\in X}(\Gamma_{AB}\Delta q^X_A+\Gamma_{BA}\Delta q^X_B)\Delta q^X_A\Delta q^X_B\ .\end{aligned} \tag{6}$$

The first three terms are exactly the same as in FMO-DFTB2 [29].

The coupling of the charge transfer to the embedding potential for dimer

IJ, ΔE^V_{IJ}, is defined as

$$\Delta E^V_{IJ} = \sum_{A\in IJ}\sum_{K\neq I,J}^{N}\sum_{C\in K}\Big\{\gamma_{AC}\Delta\Delta q^{IJ}_A \Delta q^K_C + \frac{1}{3}\Gamma_{AC}\Delta\Delta Q^{IJ}_A \Delta q^K_C + \frac{1}{3}\Gamma_{CA}\Delta\Delta q^{IJ}_A (\Delta q^K_C)^2\Big\} \ . \quad (7)$$

The charge transfer between fragments I and J for atom A is

$$\Delta\Delta q^{IJ}_A = \Delta q^{IJ}_A - \Delta q^I_A \delta_{A\in I} - \Delta q^J_A \delta_{A\in J} \ , \quad (8)$$

where $\delta_{A\in I} = 1$ if atom A belongs to fragment I, otherwise it is zero. The difference between squared charges of atoms in dimers and monomers is

$$\Delta\Delta Q^{IJ}_A = (\Delta q^{IJ}_A)^2 - (\Delta q^I_A)^2\delta_{A\in I} - (\Delta q^J_A)^2\delta_{A\in J} \ . \quad (9)$$

The latter two terms in Eq. (7) are the new terms in FMO-DFTB3, arising from the more complex form of the Coulomb interaction.

The Hamiltonian matrix of fragment X in FMO-DFTB3 is defined as

$$H^X_{\mu\nu} = H'^X_{\mu\nu} + P^X_{\mu\nu} + V^X_{\mu\nu} \ , \quad (10)$$

where $H'^X_{\mu\nu}$ is the internal Hamiltonian matrix element of fragment X (using Eq. (3) for atoms in X). $P^X_{\mu\nu}$ is the hybrid projection operator (HOP) [33], used to define fragment boundaries across covalent bonds. $V^X_{\mu\nu}$ is the electrostatic potential (ESP), describing the embedding Coulomb interaction acting on fragment X,

$$V^X_{\mu\nu} = S^X_{\mu\nu}\sum_{K\neq X}^{N}\sum_{C\in K}\Big\{\frac{1}{2}(\gamma_{AC}+\gamma_{BC}) + \frac{1}{3}(\Gamma_{AC}\Delta q^X_A + \Gamma_{BC}\Delta q^X_B) + \frac{1}{6}(\Gamma_{CA}+\Gamma_{CB})\Delta q^K_C\Big\}\Delta q^K_C \ . \quad (11)$$

In DFTB2, all Γ terms are absent; thus ESP depends only on the external charges of atoms outside X. For DFTB3, however, ESP also depends on the charges in X (the second term of Eq. (11)), and has to be updated during SCC of X. This can be efficiently accomplished by defining

$$T^X_A = \frac{2}{3}\sum_{K\neq X}^{N}\sum_{C\in K}\Gamma_{AC}\Delta q^K_C \ , \quad (12)$$

and then the DFTB3 specific ESP contribution for fragment X, which also depends on the charges in X (the second term in Eq. (11)), is calculated as $\frac{1}{2}S^X_{\mu\nu}(\Delta q^X_A T^X_A + \Delta q^X_B T^X_B)$ for $\mu \in A$ and $\nu \in B$.

In the electrostatic dimer (ES-DIM) approximation [34], the internal energy in Eq. (6) for a separated dimer IJ is evaluated as

$$E'_{IJ} \approx E'_I + E'_J + \sum_{A\in I}\sum_{B\in J}\left\{\gamma_{AB}\Delta q^I_A \Delta q^J_B + \frac{1}{3}(\Gamma_{AB}\Delta q^I_A + \Gamma_{BA}\Delta q^J_B)\Delta q^I_A \Delta q^J_B\right\} . \tag{13}$$

The repulsive energy E^{rep}_{AB} is short-ranged and for typical values of the ES-DIM thresholds (about 10 bohr) it is essentially zero, so we neglect its contribution for far separated dimers.

2.3. Analytic gradient for FMO-DFTB3

The gradients of the internal and embedding energies are needed to obtain the gradient of the total FMO-DFTB energy in Eq. (5). For the gradient derivations below, we assume a closed shell singlet state ($n_i = 2$ for all occupied MOs). The derivations below are for FMO-DFTB3, and more details on the derivations can be understood by following the derivation of the FMO-DFTB2 gradient [29]. The gradient of the internal energy in Eq. (6) is

$$\begin{aligned}
\frac{\partial E'_X}{\partial R_{\alpha x}} = \sum_{(A\neq\alpha)\in X}\sum_{\mu\in A}\sum_{\nu\in\alpha}\Bigg[&2D^X_{\mu\nu}\frac{\partial H^{0,X}_{\mu\nu}}{\partial R_{\alpha x}} - 2W'^X_{\mu\nu}\frac{\partial S^X_{\mu\nu}}{\partial R_{\alpha x}} \\
&+ D^X_{\mu\nu}\frac{\partial S^X_{\mu\nu}}{\partial R_{\alpha x}}\sum_{C\in X}\Bigg\{\gamma_{AC} + \gamma_{\alpha C} \\
&+ \frac{1}{3}(2\Gamma_{AC}\Delta q^X_A + \Gamma_{CA}\Delta q^X_C + 2\Gamma_{\alpha C}\Delta q^X_\alpha + \Gamma_{C\alpha}\Delta q^X_C)\Bigg\}\Delta q^X_C \\
&+ 2D^X_{\mu\nu}\frac{\partial P^X_{\mu\nu}}{\partial R_{\alpha x}}\Bigg] + \Delta q^X_\alpha \sum_{A\neq\alpha}\Delta q^X_A \frac{\partial \gamma_{A\alpha}}{\partial R_{\alpha x}} \\
&+ \frac{1}{3}\Delta q^X_\alpha \sum_{A\neq\alpha}\Delta q^X_A\left(\Delta q^X_A \frac{\partial \Gamma_{A\alpha}}{\partial R_{\alpha x}} + \Delta q^X_\alpha \frac{\partial \Gamma_{\alpha A}}{\partial R_{\alpha x}}\right) \\
&+ \sum_{A\neq\alpha}\frac{\partial E^{\mathrm{rep}}_{A\alpha}}{\partial R_{\alpha x}} ,
\end{aligned} \tag{14}$$

where $D^X_{\mu\nu}$ is the density matrix of fragment X and some other quantities such as $\Gamma_{\alpha C}$ are defined analogously to $\gamma_{\alpha C}$ (see [21] for details). $\tilde{W}'^X_{\mu\nu}$ is the internal Lagrangian, defined by subtracting the ESP contribution from the usual Lagrangian term $W^X_{\mu\nu}$, that is

$$\tilde{W}'^X_{\mu\nu} = W^X_{\mu\nu} - \frac{1}{2}\sum_{\rho\sigma} D^X_{\mu\rho} V^X_{\rho\sigma} D^X_{\sigma\nu} \ , \tag{15}$$

and

$$W^X_{\mu\nu} = \frac{1}{2}\sum_{\rho\sigma} D^X_{\mu\rho} H^X_{\rho\sigma} D^X_{\sigma\nu} \ . \tag{16}$$

The HOP derivative $\partial P^X_{\mu\nu}/\partial R_{\alpha x}$ is described elsewhere [35]. In this work, we did not include the residual response term contributions which require solving the complicated self-consistent Z-vector equation [36].

The derivative of the embedding energy in Eq. (7) for atom α in dimer

IJ $(\alpha \in IJ)$,

$$
\begin{aligned}
\frac{\partial \Delta E^V_{IJ}}{\partial R_{\alpha x}} &= \Delta\Delta q^{IJ}_{\alpha} \sum_{K\neq I,J}^{N} \sum_{C\in K} \Delta q^{K}_{C} \frac{\partial \gamma_{\alpha C}}{\partial R_{\alpha x}} \\
&+ \sum_{A\in IJ} \sum_{\mu\in A} \sum_{\nu\in IJ} \left(\Delta \tilde{W}^{IJ,\alpha}_{\mu\nu} S^{IJ}_{\mu\nu} + \Delta D^{IJ}_{\mu\nu} \frac{\partial S^{IJ}_{\mu\nu}}{\partial R_{\alpha x}} \right) \sum_{K\neq I,J}^{N} \sum_{C\in K} \Delta q^{K}_{C} \gamma_{\alpha C} \\
&+ \frac{1}{3} \Delta\Delta Q^{IJ}_{\alpha} \sum_{K\neq I,J}^{N} \sum_{C\in K} \Delta q^{K}_{C} \frac{\partial \Gamma_{\alpha C}}{\partial R_{\alpha x}} \\
&+ \frac{2}{3} \sum_{A\in IJ} \sum_{\mu\in A} \sum_{\nu\in IJ} \left\{ \left(\tilde{W}^{IJ,\alpha}_{\mu\nu} S^{IJ}_{\mu\nu} + D^{IJ}_{\mu\nu} \frac{\partial S^{IJ}_{\mu\nu}}{\partial R_{\alpha x}} \right) \Delta q^{IJ}_{A} \right. \\
&- \left(\tilde{W}^{I,\alpha}_{\mu\nu} S^{I}_{\mu\nu} + D^{I}_{\mu\nu} \frac{\partial S^{I}_{\mu\nu}}{\partial R_{\alpha x}} \right) \Delta q^{I}_{A} \\
&- \left. \left(\tilde{W}^{J,\alpha}_{\mu\nu} S^{J}_{\mu\nu} + D^{J}_{\mu\nu} \frac{\partial S^{J}_{\mu\nu}}{\partial R_{\alpha x}} \right) \Delta q^{J}_{A} \right\} \sum_{K\neq I,J}^{N} \sum_{C\in K} \Delta q^{K}_{C} \Gamma_{\alpha C} \\
&+ \frac{1}{3} \Delta\Delta q^{IJ}_{\alpha} \sum_{K\neq I,J}^{N} \sum_{C\in K} (\Delta q^{K}_{C})^2 \frac{\partial \Gamma_{C\alpha}}{\partial R_{\alpha x}} \\
&+ \frac{1}{3} \sum_{A\in IJ} \sum_{\mu\in A} \sum_{\nu\in IJ} \left(\Delta \tilde{W}^{IJ,\alpha}_{\mu\nu} S^{IJ}_{\mu\nu} + \Delta D^{IJ}_{\mu\nu} \frac{\partial S^{IJ}_{\mu\nu}}{\partial R_{\alpha x}} \right) \sum_{K\neq I,J}^{N} \sum_{C\in K} (\Delta q^{K}_{C})^2 \Gamma_{C\alpha},
\end{aligned}
\tag{17}
$$

and for $\alpha \in (K \neq I, J)$,

$$\begin{aligned}
\frac{\partial \Delta E^V_{IJ}}{\partial R_{\alpha x}} &= \Delta q^K_\alpha \sum_{A\in I,J} \Delta\Delta q^{IJ}_A \frac{\partial \gamma_{A\alpha}}{\partial R_{\alpha x}} \\
&+ \sum_{C\in K}\sum_{\mu\in C}\sum_{\nu\in K}\left(\tilde{W}^{K,\alpha}_{\mu\nu} S^K_{\mu\nu} + D^K_{\mu\nu}\frac{\partial S^K_{\mu\nu}}{\partial R_{\alpha x}}\right)\sum_{A\in I,J}\Delta\Delta q^{IJ}_A \gamma_{\alpha C} \\
&+ \frac{1}{3}\Delta q^K_\alpha \sum_{A\in I,J}\Delta\Delta Q^{IJ}_A \frac{\partial \Gamma_{A\alpha}}{\partial R_{\alpha x}} \\
&+ \frac{1}{3}\sum_{C\in K}\sum_{\mu\in C}\sum_{\nu\in K}\left(\tilde{W}^{K,\alpha}_{\mu\nu} S^K_{\mu\nu} + D^K_{\mu\nu}\frac{\partial S^K_{\mu\nu}}{\partial R_{\alpha x}}\right)\sum_{A\in I,J}\Delta\Delta Q^{IJ}_A \Gamma_{A\alpha} \\
&+ \frac{1}{3}(\Delta q^K_\alpha)^2 \sum_{A\in I,J}\Delta\Delta q^{IJ}_A \frac{\partial \Gamma_{\alpha A}}{\partial R_{\alpha x}} \\
&+ \frac{2}{3}\sum_{C\in K}\sum_{\mu\in C}\sum_{\nu\in K}\left(\tilde{W}^{K,\alpha}_{\mu\nu} S^K_{\mu\nu} + D^K_{\mu\nu}\frac{\partial S^K_{\mu\nu}}{\partial R_{\alpha x}}\right)\Delta q^K_C \sum_{A\in I,J}\Delta\Delta q^{IJ}_A \Gamma_{\alpha A} \ .
\end{aligned} \tag{18}$$

The definitions of $\tilde{W}^{X,\alpha}_{\mu\nu}$, $\Delta\tilde{W}^{IJ,\alpha}_{\mu\nu}$, and $\Delta D^{IJ}_{\mu\nu}$ can be found elsewhere [29].

The derivative of the ES-DIM dimer energy in Eq. (13) can be obtained

as follows, For $\alpha \in I$,

$$\begin{aligned}
\frac{\partial E'_{IJ}}{\partial R_{\alpha x}} &= \Delta q^I_\alpha \sum_{B\in J} \Delta q^J_B \frac{\partial \gamma_{\alpha B}}{\partial R_{\alpha x}} \\
&+ \sum_{A\in I}\sum_{\mu\in A}\sum_{\nu\in I}\left(\tilde{W}^{I,\alpha}_{\mu\nu} S^I_{\mu\nu} + D^I_{\mu\nu}\frac{\partial S^I_{\mu\nu}}{\partial R_{\alpha x}}\right)\sum_{B\in J}\Delta q^J_B \gamma_{\alpha B} \\
&+ \frac{1}{3}(\Delta q^I_\alpha)^2 \sum_{B\in J}\Delta q^J_B \frac{\partial \Gamma_{\alpha B}}{\partial R_{\alpha x}} \\
&+ \frac{2}{3}\sum_{A\in I}\sum_{\mu\in A}\sum_{\nu\in I}\left(\tilde{W}^{I,\alpha}_{\mu\nu} S^I_{\mu\nu} + D^I_{\mu\nu}\frac{\partial S^I_{\mu\nu}}{\partial R_{\alpha x}}\right)\Delta q^I_A \sum_{B\in J}\Delta q^J_B \Gamma_{\alpha B} \\
&+ \frac{1}{3}\Delta q^I_\alpha \sum_{B\in J}(\Delta q^J_B)^2 \frac{\partial \Gamma_{B\alpha}}{\partial R_{\alpha x}} \\
&+ \frac{1}{3}\sum_{A\in I}\sum_{\mu\in A}\sum_{\nu\in I}\left(\tilde{W}^{I,\alpha}_{\mu\nu} S^I_{\mu\nu} + D^I_{\mu\nu}\frac{\partial S^I_{\mu\nu}}{\partial R_{\alpha x}}\right)\sum_{B\in J}(\Delta q^J_B)^2 \Gamma_{B\alpha}\ ,
\end{aligned} \tag{19}$$

and for $\alpha \in J$,

$$\begin{aligned}
\frac{\partial E'_{IJ}}{\partial R_{\alpha x}} &= \Delta q^J_\alpha \sum_{A\in I} \Delta q^I_A \frac{\partial \gamma_{A\alpha}}{\partial R_{\alpha x}} \\
&+ \sum_{B\in J}\sum_{\mu\in B}\sum_{\nu\in J}\left(\tilde{W}^{J,\alpha}_{\mu\nu} S^J_{\mu\nu} + D^J_{\mu\nu}\frac{\partial S^J_{\mu\nu}}{\partial R_{\alpha x}}\right)\sum_{A\in I}\Delta q^I_A \gamma_{A\alpha} \\
&+ \frac{1}{3}\Delta q^J_\alpha \sum_{A\in I}(\Delta q^I_A)^2 \frac{\partial \Gamma_{A\alpha}}{\partial R_{\alpha x}} \\
&+ \frac{1}{3}\sum_{B\in J}\sum_{\mu\in B}\sum_{\nu\in J}\left(\tilde{W}^{J,\alpha}_{\mu\nu} S^J_{\mu\nu} + D^J_{\mu\nu}\frac{\partial S^J_{\mu\nu}}{\partial R_{\alpha x}}\right)\sum_{A\in I}(\Delta q^I_A)^2 \Gamma_{A\alpha} \\
&+ \frac{1}{3}(\Delta q^J_\alpha)^2 \sum_{A\in I}\Delta q^I_A \frac{\partial \Gamma_{\alpha A}}{\partial R_{\alpha x}} \\
&+ \frac{2}{3}\sum_{B\in J}\sum_{\mu\in B}\sum_{\nu\in J}\left(\tilde{W}^{J,\alpha}_{\mu\nu} S^J_{\mu\nu} + D^J_{\mu\nu}\frac{\partial S^J_{\mu\nu}}{\partial R_{\alpha x}}\right)\Delta q^J_B \sum_{A\in I}\Delta q^I_A \Gamma_{\alpha A}\ .
\end{aligned} \tag{20}$$

2.4. Damping of the γ function

The γ function appearing in DFTB2 and DFTB3 describes the distance dependence of the Coulomb interaction of two point charges [37]. It can be

written as

$$\gamma_{AB} = \frac{1}{r_{AB}} - S(r_{AB}, U_A, U_B) \ , \tag{21}$$

where U_A is the Hubbard parameter, depending on the chemical hardness of atom A [30]. S is a short-range function [21, 37] enforcing the correct behavior of γ_{AB} at $r_{AB} = 0$. Ideally, the γ function should approach U_A at a short distance, but it is not the case for H atom due to an inconsistency between the atomic size and the chemical hardness [19, 30]. The consequence is the underestimation of the binding energy of hydrogen-bonded complexes by about 1-2 kcal/mol per bond [30]. This problem can be removed by damping the γ function at short distances,

$$\gamma^{\mathrm{h}}_{AB} = \frac{1}{r_{AB}} - S(r_{AB}, U_A, U_B) h(r_{AB}, U_A, U_B) \ , \tag{22}$$

where $h(r_{AB}, U_A, U_B) = 1$ if neither A nor B is a H atom, and otherwise

$$h(r_{AB}, U_A, U_B) = \exp\left\{ -\left(\frac{U_A + U_B}{2} \right)^{\zeta} r^2_{AB} \right\} \ , \tag{23}$$

where ζ is a parameter. In this work, the damped function γ^{h} is used instead of γ in all DFTB3 calculations.

3. Computational details

The DFTB3 and FMO-DFTB3 methods were implemented in GAMESS-US [38] and the latter was parallelized with the generalized distributed data interface (GDDI) [39]. In the earlier FMO-DFTB2 code [29], we used numeric gradients of the non-perturbed Hamiltonian $H^0_{\mu\nu}$ and overlap $S_{\mu\nu}$ matrix elements; in this work we implemented their analytic derivatives.

The 3ob set of DFTB parameters with the optimized Hubbard derivatives and exponent of γ^{h} [22] were used in all DFTB calculations [40]. The unitless threshold for the ES-DIM approximation was set to 2.0. Geometry optimizations were performed until the root mean square (RMS) and maximum gradient values became smaller than $1/3 \times 10^{-4}$ and 10^{-4} hartree/bohr, respectively (OPTTOL=10^{-4} in GAMESS). Hybrid orbitals for HOP operators were generated using Pipek-Mezey orbital localization [41] and the mio parameter set [37] for sp^3 carbon by calculating methane.

The accuracy of FMO-DFTB3 is evaluated for the α-helix, β-turn, and extended form of the polyalanine COMe-$(Ala)_{20}$-NHMe, whose initial structures were optimized with TINKER 6.0 [42] using the AMBER99 force field parameters [43]. The fragmentation of the polyalanines was performed at the C_α atoms as usual.

To demonstrate the potential of FMO-DFTB, we optimized a nano flake of cellulose Iβ using FMO-DFTB3 with the Slater-Kirkwood [44] (SK) and UFF [45, 46] type dispersion corrections (denoted with the suffix "-D(SK)" and "-D(UFF)" respectively). The initial structure of the cellulose was extracted from the X-ray experimental data [47], by taking 64 chains arranged in 8 sheets, with the total of 512 glucose residues (10,944 atoms in total). The fragmentation of the cellulose was performed between a 6-membered ring and the adjacent oxygen bridge [48] with one glucose residue per fragment, using the default fragmentation of saccharides in Facio [49].

4. Results and discussion

To check the accuracy of FMO-DFTB3 compared to full (i.e., without fragmentation) DFTB3, we performed geometry optimizations of COMe-$(Ala)_{20}$-NHMe, which in FMO was divided by assigning 1, 2, 4, and 5 residues per fragment, so that there were 20, 10, 5 and 4 fragments, respectively. Table 1 shows the relative energies and RMS deviation (RMSD) for the optimized structures of the α-helix, β-turn, and extended form. Hydrogen atoms are included and mass-weighting is not used in calculating RMSD.

Table 1: The energies and RMSD (Å) between the structures of COMe-$(Ala)_{20}$-NHMe optimized with FMO-DFTB3 and full DFTB3

		FMO-DFTB3	FMO-DFTB3	FMO-DFTB3	FMO-DFTB3	full DFTB3
	fragments	20	10	5	4	1
energy[a]	α-helix	-268.340035	-268.340985	-268.341122	-268.341253	-268.341364
	β-turn	-268.311062 (+18.2)	-268.313574 (+17.2)	-268.314725 (+16.6)	-268.315069 (+16.4)	-268.315588 (+16.2)
	extended	-268.258793 (+51.0)	-268.258964 (+51.5)	-268.258988 (+51.5)	-268.258987 (+51.6)	-268.258958 (+51.7)
RMSD[b]	α-helix	0.398 (0.035)	0.037 (0.042)	0.008 (0.028)	0.008 (0.021)	0.000
	β-turn	0.393 (0.067)	0.064 (0.026)	0.037 (0.032)	0.019 (0.044)	0.000
	extended	0.159 (0.045)	0.203 (0.054)	0.193 (0.000)	0.195 (0.000)	0.000

[a] The total energies are in hartree, and the values in parentheses are the energies relative to the α-helix (kcal/mol). [b] RMSD values are shown for the FMO-DFTB3 and DFTB3 optimizations starting from the same initial MM-optimized geometry; the values in parentheses are for DFTB3 optimizations starting from the FMO-DFTB3 optimized structures in an attempt to eliminate a difference in local minima.

Full DFTB3 predicts that the optimized β-turn and extended structures are less stable than the α-helix by 16.2 and 51.7 kcal/mol, respectively (full DFTB2 values [29] are 18.6 and 47.8 kcal/mol, respectively). The error in the energy for FMO-DFTB3 optimized structures compared to full DFTB3 is at most 2.0 kcal/mol for 1 residue per fragment, while for larger fragment sizes the error becomes 1 kcal/mol or less. FMO-DFTB3 errors reported here are similar to those of FMO-DFTB2 [29].

Polyalanines are flexible with a flat energy surface featuring many local minima. The geometry optimized with MM is relatively far from the DFTB3 minima, and the minima obtained with FMO-DFTB3 and full DFTB3 starting from the MM minima show RMSD values of 0.4 Å for 1 residue per fragment; for larger fragment sizes RMSD does not exceed 0.064 Å for the more rigid α-helix and β-turn, while for the more flexible extended form the values are about 0.2 Å or less. However, the deviations become much smaller if the FMO-DFTB3 optimized geometry was used as the initial structure in the full DFTB3 optimization, in which case the largest deviation of FMO-DFTB3 from full DFTB3 is 0.067 Å for 1 residue per fragment, and smaller for larger fragment sizes.

The numeric contribution of neglected residual response terms was estimated for the α-helix of polyalanine with 1 and 5 residues per fragment, by calculating the difference between the analytic and numeric gradients. The RMS deviations were 0.000 153 and 0.000 033 hartree/bohr for 1 and 5 residues per fragment, respectively. The RMS deviation between analytic FMO-DFTB3 and full DFTB3 gradients was 0.000 230 and 0.000 039 hartree/bohr for 1 and 5 residues per fragment, respectively.

For an application of FMO-DFTB3, we optimized a cluster of cellulose Iβ consisting of 10,944 atoms (Figure 1), starting from the experimental X-ray structure. Because this is a finite-size cluster (a flake), there are some interesting edge effects as shown in Figure 2. In order to compare the geometry of the bulk of the optimized cluster with the experimental structure of a macro-sized cellulose, we focus on the central part of the optimized cluster, defined in two ways: (a) the central cluster of about 4,500 atoms and (b) the middle sheet in (a), consisting of 762 atoms, called central sheet below. When computing RMSD, H atoms were excluded because of comparing to X-ray experiment.

Table 2 shows that the RMSD values do not exceed 0.47 Å for the central cluster. The central cluster containing several sheets includes the effect of sheet interactions, and in order to quantify this effect and separate it from

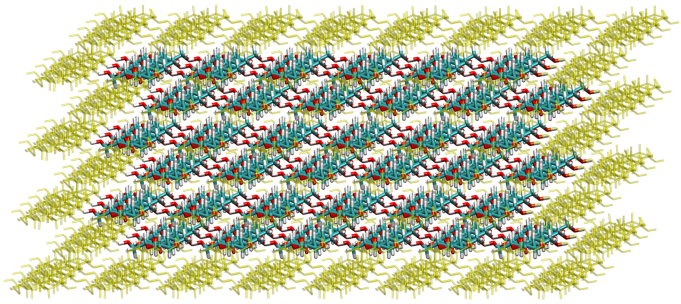

Figure 1: Cellulose Iβ model (initial structure). For the central cluster, hydrogen, carbon, and oxygen atoms are shown in white, cyan, and red, respectively; other atoms (not included in computing RMSD to experiment) are shown in yellow.

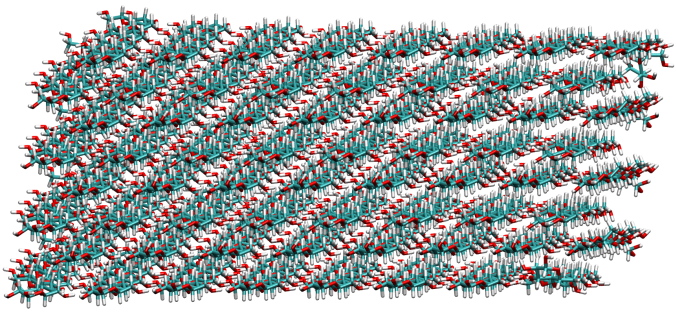

Figure 2: Optimized structure of cellulose Iβ with FMO-DFTB3-D(SK). Hydrogen, carbon, and oxygen atoms are shown in white, cyan, and red, respectively.

Table 2: RMSD (Å) between optimized (FMO-DFTB3-D) and experimental [47] structures of cellulose Iβ and the FMO-DFTB3-D binding energies (kcal/mol) per glucose residue between two sheets ($\Delta E^{\mathrm{sheet}}$) and two chains ($\Delta E^{\mathrm{chain}}$).

	central cluster[a]		central sheet[a]	
dispersion	RMSD	$\Delta E^{\mathrm{sheet}}$	RMSD	$\Delta E^{\mathrm{chain}}$
SK	0.467	9.34	0.232	8.30
UFF	0.465	12.45	0.351	7.63

[a] See Figure 1 for the definition of the central cluster in the 10,944 atomic system optimized with FMO-DFTB3, and the central sheet is the middle sheet in it.

the role of chain interactions within a sheet, we calculated RMSD for the third (middle) sheet of the central cluster from the top (the fourth sheet in the optimized structure).

The RMSD for the sheet is smaller than for the cluster because of the packing effects affecting the interaction between sheets, which are more strongly affected by the truncation of the macro-sized cellulose into a nano flake than the interactions within a sheet. Comparing the two dispersion models, the RMSDs for the central cluster are very small, but for the central sheet the RMSD with the Slater-Kirkwood has a much smaller deviation from experiment (0.232 Å compared to 0.351 Å with the UFF model).

The intra-sheet interaction is dominated by conventional hydrogen bonding interaction. On the other hand, the inter-sheet interaction can be described as "non-conventional" hydrogen bonding [31] dominated by dispersion. The inter-sheet distances for X-ray, FMO-DFTB3-D(SK) and FMO-DFTB3-D(UFF) optimized structures are 7.78, 7.38, and 7.51 Å, respectively, which means that in the direction perpendicular to the sheet plane, DFTB3 predicts a small shrinking of the nano flake compared to bulk cellulose. The inter-chain distances (i.e., intra-sheet distances along hydrogen bonds) for X-ray, FMO-DFTB3-D(SK) and FMO-DFTB3-D(UFF) are 8.20, 8.15, and 8.30 Å, respectively, so FMO-DFTB3-D(UFF) shows small expansion by 0.10 Å, while FMO-DFTB3-D(SK) results in a shrinking by 0.05 Å.

The inter-sheet distances imply that the Slater-Kirkwood dispersion shows a stronger inter-sheet interaction than the UFF-type model. Although the intra-sheet interaction which comes from conventional hydrogen bonds is well described with the Slater-Kirkwood dispersion (the smaller RMSD for the central sheet), it overestimated the inter-sheet interaction. UFF does the opposite; it underestimates conventional hydrogen bonding, while inter-sheet interaction is better than the Slater-Kirkwood dispersion correction. The net result is that the two dispersion models show very similar RMSDs for the central cluster.

The interactions can be quantitatively analyzed by comparing the binding energies per glucose residue between sheets ($\Delta E^{\mathrm{sheet}}$) and chains ($\Delta E^{\mathrm{chain}}$),

$$\Delta E^{\mathrm{sheet}} = \frac{1}{64}(E^{\mathrm{sheet}} - \frac{1}{8}E^{\mathrm{flake}}) \ , \tag{24}$$

and

$$\Delta E^{\mathrm{chain}} = \frac{1}{8}(E^{\mathrm{chain}} - \frac{1}{8}E^{\mathrm{sheet}}) \ , \tag{25}$$

where E^{flake}, E^{sheet}, and E^{chain} are the total energies of the structures optimized with FMO-DFTB3 for the total flake (10,944 atoms), a single sheet (1,368 atoms), and a chain (171 atoms), respectively. The fractions in Eqs. (24) and (25) represent the number of glucose residues and other subunits. The binding energies shown in Table 2 are consistent with the above discussion based on the structural aspects only; ΔE^{chain} for FMO-DFTB3-D(SK) is bigger than that for FMO-DFTB3-D(UFF), and the opposite is observed for ΔE^{sheet}.

In order to estimate how much the results are affected by the fragment size, we did the optimization of the cellulose flake using the fragmentation of 2 glucose residues per fragment, starting from the same initial structure. The RMSD of 0.467 Å in case of FMO-DFTB3-D(SK) for the central cluster (1 glucose per fragment) is improved to 0.453 Å with two glucose residues per fragment. The inter-sheet binding energy per sheet is essentially not affected: it changes by 0.01 kcal/mol to 9.35 kcal/mol (cf. Table 2) with larger fragments. Therefore, it can be concluded that FMO-DFTB3 accurately predicts the full DFTB3 structure of cellulose and the difference of the optimized structure to experiment comes from DFTB3 deficiencies and from the difference between a flake and bulk cellulose.

The intra-sheet binding energies between two chains obtained with DFTB method are compared with spin-component scaled second-order Møller-Plesset perturbation theory with the resolution of the identity (RI-SCS-MP2, performed using TURBOMOLE[50]). Table 3 shows the binding energies at different levels of theory for a small cluster optimized at full DFTB3-D(SK) level of theory (Figure 3). Comparing DFTB3 with RI-SCS-MP2/aug-cc-pVTZ, the difference in the binding energy is 1.44 and 2.49 kcal/mol for the D(SK) and D(UFF) models, respectively, so the former is better for cellulose.

The error introduced by the FMO method for the system in Table 3 was about 0.15 kcal/mol. Note that the dispersion correction in FMO-DFTB exactly reproduces that in full DFTB, because the dispersion energy is independent from the electronic structure (it is a parameterized *ad hoc* correction).

The wall-clock timings for one gradient calculation of a cellulose sheet in the flake (1,368 atoms) and the whole flake (10,944 atoms) with FMO-DFTB3-D(SK) using 16 CPU cores (1 node) of dual E5-2650 Xeon CPUs were 2.3 and 72.4 seconds. For comparison, the same cellulose sheet calculated on 1 CPU core of E5-2650 Xeon took for full DFTB3 and FMO-DFTB3 1148.9 and 26.4 seconds, respectively, i.e., FMO makes the calculation 43.5 times

Table 3: Binding energies (kcal/mol) per glucose residue calculated for the structure shown in Figure 3 optimized with DFTB3-D(SK).

method	binding energy
DFTB3-D(SK)	9.47
DFTB3-D(UFF)	8.42
RI-MP2/aug-cc-pVTZ	12.36
RI-SCS-MP2/aug-cc-pVTZ	10.91

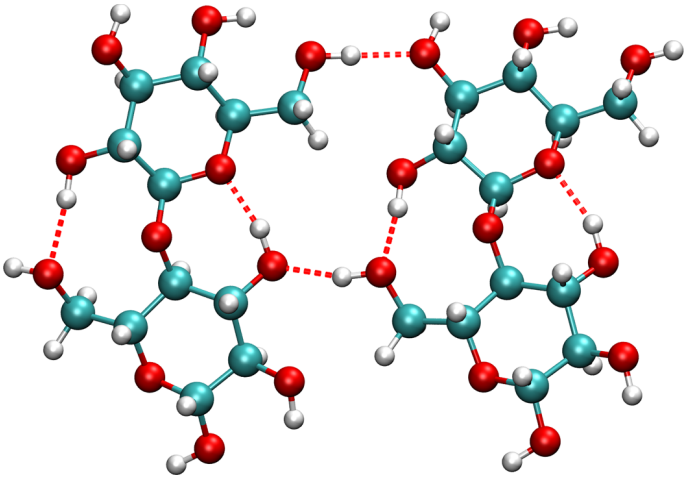

Figure 3: Structure of the small cellulose Iβ fragment optimized with full DFTB3-D(SK). Hydrogen bonds are shown as dashed red lines.

faster. Full DFTB3-D(SK) single point calculation for the optimized sheet structure with FMO-DFTB3-D(SK) gives the difference of 10.60 kcal/mol in total energy. The error is reduced to 4.63 kcal/mol by doubling the fragment size of FMO-DFTB3-D(SK) single point calculation with the same structure.

5. Conclusions

We have extended the FMO-DFTB method to include the third-order correction (FMO-DFTB3) and implemented the damping of the γ function. FMO-DFTB3 has been developed in GAMESS-US and parallelized using GDDI. FMO-DFTB3 has been shown to be consistently accurate for three isomers of polyalanine, α-helix, β-turn, and the extended form, both for the energy and the optimized structures. Using FMO, we accelerated DFTB3 calculations for a cellulose sheet containing 1,368 atoms by the factor of 43.5.

By optimizing a nano flake of cellulose Iβ (10,944 atoms), we have demonstrated the usefulness of FMO-DFTB3 and its accuracy in comparison to the experimental results. Two dispersion models have been compared and their advantages and drawbacks for cellulose have been highlighted. We hope that FMO-DFTB3 will be useful in future applications to large molecular systems.

Acknowledgements

The authors thank the Research Center for Computational Science, Okazaki, Japan for providing computational resources. Y. N. is supported by a Research Fellowships of the Japan Society for Promotion of Science for Young Scientists (DC1), and the L-Daigakuin program. D. G. F. acknowledges support of the Next Generation Super Computing Project, Nanoscience Program (MEXT, Japan) and D. G. F. and S. I. are grateful for support from the Computational Materials Science Initiative (CMSI, Japan).

References

[1] W. Foulkes, L. Mitas, R. Needs, G. Rajagopal, Rev. Mod. Phys. 73 (2001) 33–83.

[2] B. M. Austin, D. Y. Zubarev, W. A. Lester, Chem. Rev. 112 (2012) 263–288.

[3] G. P. Szalay, T. Müller, G. Gidofalvi, H. Lischka, R. Shepard, Chem. Rev. 112 (2012) 108–181.

[4] D. I. Lyakh, M. Musiał, V. F. Lotrich, R. J. Bartlett, Chem. Rev. 112 (2012) 182–243.

[5] Y. Andoh, N. Yoshii, K. Fujimoto, K. Mizutani, H. Kojima, A. Yamada, S. Okazaki, K. Kawaguchi, H. Nagao, K. Iwahashi, F. Mizutani, K. Minami, S.-I. Ichikawa, H. Komatsu, S. Ishizuki, Y. Takeda, M. Fukushima, J. Chem. Theory Comput. 9 (2013) 3201–3209.

[6] S. Goedecker, G. E. Scuseria, Comp. Sci. Eng. 5 (2003) 14–21.

[7] M. S. Gordon, D. G. Fedorov, S. R. Pruitt, L. V. Slipchenko, Chem. Rev. 112 (2012) 632–672.

[8] P. Otto, J. Ladik, Chem. Phys. 8 (1975) 192–200.

[9] W. Yang, Phys. Rev. Lett. 66 (1991) 1438–1441.

[10] J. Gao, J. Phys. Chem. B 101 (1997) 657–663.

[11] H. Yu, H. R. Leverentz, P. Bai, J. I. Siepmann, D. G. Truhlar, J. Phys. Chem. Lett. 5 (2014) 660–670.

[12] M. A. Collins, Phys. Chem. Chem. Phys. 14 (2012) 7744–7751.

[13] K. Wang, W. Li, S. Li, J. Chem. Theory Comput. 10 (2014) 1546–1553.

[14] N. Sahu, S. D. Yeole, S. R. Gadre, J. Chem. Phys. 138 (2013) 104101.

[15] P. Koskinen, V. Mäkinen, Comput. Mater. Sci. 47 (2009) 237–253.

[16] A. F. Oliveira, G. Seifert, T. Heine, H. A. Duarte, J. Braz. Chem. Soc. 20 (2009) 1193–1205.

[17] G. Seifert, J.-O. Joswig, WIREs Comput. Mol. Sci. 2 (2014) 456–465.

[18] M. Gaus, Q. Cui, M. Elstner, WIREs Comput. Mol. Sci. 4 (2014) 49–61.

[19] M. Elstner, Theor. Chem. Acc. 116 (2006) 316–325.

[20] Y. Yang, H. Yu, D. York, Q. Cui, M. Elstner, J. Phys. Chem. A 111 (2007) 10861–10873.

[21] M. Gaus, Q. Cui, M. Elstner, J. Chem. Theory Comput. 7 (2011) 931–948.

[22] M. Gaus, A. Goez, M. Elstner, J. Chem. Theory Comput. 9 (2013) 338–354.

[23] P. Goyal, M. Elstner, Q. Cui, J. Phys. Chem. B 115 (2011) 6790–6805.

[24] P. Goyal, H.-J. Qian, S. Irle, X. Lu, D. Roston, T. Mori, M. Elstner, Q. Cui, J. Phys. Chem. B 118 (2014) 11007–11027.

[25] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, Chem. Phys. Lett. 313 (1999) 701–706.

[26] D. G. Fedorov, K. Kitaura, J. Phys. Chem. A. 111 (2007) 6904–6914.

[27] D. G. Fedorov, T. Nagata, K. Kitaura, Phys. Chem. Chem. Phys. 14 (2012) 7562–7577.

[28] S. Tanaka, Y. Mochizuki, Y. Komeiji, Y. Okiyama, K. Fukuzawa, Phys. Chem. Chem. Phys. 16 (2014) 10310–10344.

[29] Y. Nishimoto, D. G. Fedorov, S. Irle, J. Chem. Theory Comput. 10 (2014) 4801–4812.

[30] M. Elstner, J. Phys. Chem. A 111 (2007) 5614–5621.

[31] A. Devarajan, S. Markutsya, M. H. Lamm, X. Cheng, J. C. Smith, J. Y. Baluyut, Y. Kholod, M. S. Gordon, T. L. Windus, J. Phys. Chem. B 117 (2013) 10430–10443.

[32] R. G. Parr, R. G. Pearson, J. Am. Chem. Soc. 105 (1983) 7512–7516.

[33] T. Nakano, T. Kaminuma, T. Sato, Y. Akiyama, M. Uebayasi, K. Kitaura, Chem. Phys. Lett. 318 (2000) 614–618.

[34] T. Nakano, T. Kaminuma, T. Sato, K. Fukuzawa, Y. Akiyama, M. Uebayasi, K. Kitaura, Chem. Phys. Lett. 351 (2002) 475–480.

[35] T. Nagata, D. G. Fedorov, K. Kitaura, Chem. Phys. Lett. 492 (2010) 302–308.

[36] T. Nagata, K. Brorsen, D. G. Fedorov, K. Kitaura, M. S. Gordon, J. Chem. Phys. 134 (2011) 124115.

[37] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, G. Seifert, Phys. Rev. B 58 (1998) 7260–7268.

[38] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. J. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, J. Comput. Chem. 14 (1993) 1347–1363.

[39] D. G. Fedorov, R. M. Olson, K. Kitaura, M. S. Gordon, S. Koseki, J. Comput. Chem. 25 (2004) 872–880.

[40] The dftb web site, http://www.dftb.org .

[41] J. Pipek, P. G. Mezey, J. Chem. Phys. 90 (1989) 4916–4926.

[42] J. W. Ponder, TINKER: Software Tools for Molecular Design, 6.0, Washington University School of Medicine, Saint Louis, MO (2009).

[43] J. Wang, P. Cieplak, P. A. Kollman, J. Comput. Chem. 21 (2000) 1049–1074.

[44] M. Elstner, P. Hobza, T. Frauenheim, S. Suhai, E. Kaxiras, J. Chem. Phys. 114 (2001) 5149–5155.

[45] A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, W. M. Skiff, J. Am. Chem. Soc. 114 (1992) 10024–10035.

[46] L. Zhechkov, T. Heine, S. Patchkovskii, G. Seifert, H. A. Duarte, J. Chem. Theory Comput. 1 (2005) 841–847.

[47] Y. Nishiyama, P. Langan, H. Chanzy, J. Am. Soc. Chem. 124 (2002) 9074–9082.

[48] T. Sawada, D. G. Fedorov, K. Kitaura, Int. J. Quant. Chem. 109 (2002) 2033–2045.

[49] M. Suenaga, J. Comput. Chem. Jpn. 4 (2005) 25–32.

[50] TURBOMOLE V6.6 2014, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from
`http://www.turbomole.com`.